Communication

Composing RNA nanostructures from a syntax of RNA structural modules

Cody Geary, Arkadiusz Chworos, Erik Verzemnieks, Neil R. Voss, and Luc Jaeger

Nano Lett., **Just Accepted Manuscript** • DOI: 10.1021/acs.nanolett.7b03842 • Publication Date (Web): 17 Oct 2017

Composing RNA nanostructures from a syntax of RNA structural modules

Cody Geary[1†], Arkadiusz Chworos[1†], Erik Verzemnieks[1†], Neil R. Voss[2] and Luc Jaeger[1*]

1, Department of Chemistry and Biochemistry, Biomolecular Science and Engineering program, University of California at Santa Barbara, Santa Barbara, CA 93106-9510.

2, Biological, Chemical and Physical Sciences Department, Roosevelt University, 1400 N. Roosevelt Blvd, Schaumburg, IL 60173.

† Present address: Bioengineering, Computer Science, and Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125, USA (C.G.). Center of Molecular and Macromolecular Studies of Polish Academy of Sciences, 90-363 Lodz, Poland (A.C); Swedish Medical Center, Emergency Department, 501E Hampden Ave, Englewood, CO 80113 (E.V.).

*Correspondence to: jaeger@chem.ucsb.edu: Phone: 805 893 3628; Fax: 805 893 4120

Abstract

Natural stable RNAs fold and assemble into complex three-dimensional architectures by relying on the hierarchical formation of intricate, recurrent networks of non-covalent tertiary interactions. These sequence-dependent networks specify RNA structural modules enabling orientational and topological control of helical struts to form larger self-folding domains. Borrowing concepts from linguistics, we defined an extended structural syntax of RNA modules for programming RNA strands to assemble into complex, responsive nanostructures under both thermodynamic and kinetic control. Based on this syntax, various RNA building blocks promote the multi-molecular assembly of objects with well-defined three-dimensional shapes as well as the isothermal folding of long RNAs into complex single stranded nanostructures during transcription. This work offers a glimpse of the limitless potential of RNA as an informational medium for designing programmable and functional nano-materials useful for synthetic biology, nanomedicine and nanotechnology.

Key words: RNA nanotechnology, RNA self-assembly, RNA folding, nanostructures, nano-particles, tectoRNAs, RNA architectonics

The rational design and engineering of complex three-dimensional (3D) nanostructures of RNA remain a great challenge for nanotechnology, nano-chemistry, synthetic biology and nanomedicine[1-8]. In nature, functional RNAs like riboswitches and ribozymes, fold into 3D shapes through the sequence-dependent formation of exquisite 3D networks of non-covalent interactions. While the RNA secondary (2D) structure is mostly driven by the formation of local RNA helices, the RNA 3D structure is induced by recurrent tertiary modules[9-11] that constitute the lexicon enabling control of the RNA folding and assembly into larger self-folding domains[12] through the orientation, stacking and bending of adjacent helical struts, and the formation of long-range interactions (Figure 1 and supplementary Figure S1). We previously developed the RNA architectonics approach,[5, 13] a method that aids designing and building self-assembling nano-architectures from the structural information of natural RNAs.[14-20] So far, RNA nanodesign has relied on the repetitious use of a very limited repertoire of structural modules, short RNA sequences promoting self-assembly through Watson-Crick base pairs,[6, 15, 17-19] and strategies that seldom take advantage of the ability of RNA to fold and assemble during transcription in isothermal conditions[17, 21, 22]. Borrowing concepts from linguistics, we hypothesized that we could describe and generate complex RNA structures from an extended syntax: RNA modules act as 'words' that are arranged with respect to one another to define larger self-assembling domains (or tectoRNA units), thereby expressing nanostructures with complex and well-defined 3D shapes specified by the context of the constituent modules – reminiscent of the way that words make up sentences which express meanings. We demonstrate herein that such a syntax of RNA modules offers a limitless potential of expression for building complex, responsive 3D nanostructures under thermodynamic and kinetic control (Figure 1 and Supplementary Figures S1, S2).

RNA structural modules (Figure 1, left) define a wide range of possible geometries that can be used to join helical struts. The UA_handle three-way junction (UAh_3WJ) and class II tRNA (tRNA-5WJ) are 90° bend modules suitable for the construction of tetrameric assemblies, also called tectosquares.[10, 18, 23] While both modules can flex, the UAh_3WJ module is more planar and one of its helical stems preferentially flexes in one direction with respect to the two others, leading to straight, obtuse and acute angles.[24] Obtuse and acute units were both programmed with kissing-loops (KLs) to either assemble into tectosquares or tectotriangles (Figures 1, S3). Tectotriangles TT with acute units are more thermostable than those with obtuse units (TTo). The opposite trend is observed for tectosquares (Figures 2A, S4). Tectotriangles TT are nonetheless more thermostable than tectosquares TS with obtuse units. By varying the length of their arms bearing complementary tail-tail ends,

tectosquares TS and tectotriangles TT can be controlled to assemble into regular nano-arrays, ladders and nano-grids (Figures 1, 2B, S5A-C, Table S1AB). Because of the preferential, directional flexing of the UA_h-3WJ module, obtuse units designed for tectotriangles TTo also assemble into tectohexagons (Figures 2, S4). When programmed with tail-tail interactions, they form mixed TTo-2TTo "sponge-like" networks (Figures 2B, S5A). These results demonstrate that UAh_3WJ units can serve different geometrical purposes in a programmable and context dependent fashion.

The A-minor-3WJ junction module (Figure S1) controls the preferential stacking of adjacent RNA helices according to its thermodynamic stability.[19] With magnesium and increasing RNA concentrations, A-minor-3WJ units TF1 assemble end-to-end through kissing-loops into nanorings and nanofibers of various sizes, whereas units TF2, with a thermostable GAAA/11nt-3WJ junction, form long nanofibers[19] (Figures 1, S6, S7, Table S1C). AMP-TF1 and AMP-TF2 were derived from these tectoRNAs by incorporating the AMP aptamer[25-27] into their design (Figures 1, S1, S6). The AMP aptamer offers structural modulation of the flexibility of RNA bending: upon recognition of AMP (adenosine or ATP),[25, 26] AMP-TF units form a ~120° bend by induced-fit folding. Consequently, AMP-TF1 and AMP-TF2 nanostructures can be controlled in response to ligand binding. Native PAGE analysis indicates that AMP-TF1 and AMP-TF2 assemble into mixtures of tetramers, hexamers and nanofibers in presence of AMP whereas they mostly form dimers in absence of the cognate ligand (Figures 2D, S7). AMP-TF1 and AMP-TF2 variants, which inactivate the AMP-aptamer, behave as in absence of AMP (Figures 2D, S7). AFM imaging reveals that AMP-TF1 units generate long dynamical nanofibers with numerous kinks that suggest a spiraling curvature induced by AMP (Figure S7F).

To expand the versatility of these RNA modules, new tectoRNA units were derived from those above to create responsive nanostructures with distinctive shapes. Heart shapes were chosen because they require a complex collection of different bends and coaxial stacks; they are planar and easily identifiable by microscopy. As for their universal meaning, there is no greater message than "LOVE" to give to the world. We first designed tectoRNA units able to self-assemble into G1 and G2 nano-hearts (Figures 1, 3). In a second step, these units acting as structural domains were encoded into a single stranded RNA to fold into G3 nano-hearts (Figure 4).

The design of the G1-heart minimizes the number of different RNA units necessary to build a symmetrical six-unit object assembling through four distinct kissing-loops. Units HX and HY, which are respectively derived from TF and AMP-TF units, are used twice to form the two sides of the heart. They are joined by HW and HZ, two 90° angle units derived from tRNA-5WJ units that were

previously used in tectosquares and polyhedrons[18, 23] (Figures 1, S6, S8A; Table S1D). Consequently, partial control of the positioning of HW with respect of HZ results in two distinct cis and trans assembly products (Figure S8A). Unit HX forms the straight edges of the heart: alongside an A-minor-3WJ module, it incorporates a small A-bulge module that induces a slight 160° curvature to the sides of the object and also relieves the heart shape structure from possible structural stress (Figures 1, S1). Unit HY, which forms the two lobes of the heart, merges features from AMP-TF1 and AMP-TF2 units. Its strong curvature is induced by two consecutive AMP aptamers connected through an A-minor-3WJ module. The G1-heart assembles with ~70% yields on native PAGE (Figure S9). Structural Pb(II) probing of unit HY, alone or within the context of G1-hearts, indicates that its AMP-aptamer modules fold into compact, rigid bends in both contexts upon addition of AMP (Figures 3, S10). Several discreet particles can unambiguously be identified as nano-hearts by AFM (Figures 3, S11). However, the protruding tRNA-5WJ modules within the heterogeneous mix of cis and trans G1-hearts do not favor RNA adhesion on the mica surface under solution, rendering AFM visualization challenging at high resolution (Figures 3, S11).

The G2-heart, derived from the G1-heart, is designed to resemble a "pierced" heart. As an asymmetric and addressable object, its six RNA units are easily identifiable by microscopy (Figures 1, S6, S8B; Table S1D). Units HA and HD, which replace units HZ and HW from the G1-heart, are derived from TT and TS units, respectively (Figures 1, S6). Their constituent AUh_3WJ module confers an increased flatness compared to other 90° modules[15, 16, 23, 28]. Units HC and HE are structurally identical to unit HY but programmed with different kissing-loops. Units HB and HF are both derived from HX: instead of an A-minor-3WJ, HB has a UA_h_3WJ module whereas HF contains the GAAA/11nt-3WJ of TF2 (Figures 1, S6). HB forms the G2-heart crossbar by assembling with unit HD through a tail-tail triple interaction (Figure S6C). The G2-heart design presents the UA_h_3WJ module in three different orientations, leading to ~65°, 115° and 180° angles. In presence of AMP, the G2-heart assembles with a yield of ~85% as determined by native PAGE and cryo-electron-microscopy (cryo-EM) (Figures S9, S13). Approximately ~60% of the objects visualized by AFM are single G2-hearts, out of which 95% present a well-defined crossbar (Figures 3, S12). About ~75% of them present the crossbar on the right, suggesting that the overall 3D shape of G2-hearts leads to preferential orientation on the mica surface (Figure S12). This bias is not observed by cryo-EM (Figures 3, S13D). In high-resolution AFM images, the heart shape and dimension closely match the 3D model, with regions of increased height corresponding to the locations of tertiary modules within

the assembly (Figures 3, S13C). Likewise, a particle average of 2,456 particles visualized by cryo-EM shows increased density only in areas corresponding to tertiary modules, and corroborates their orientational and topological properties (Figure S13D-F). Notably, the three different orientations of the UAh_3WJ module in units HA, HB and HD, are validated by AFM and cryoEM data (Figure S13). The ability to control precise helical-stacking through a single tail-tail interaction of 7 base pairs was also tested for G2-hearts. Two sets of hearts were designed and assembled through a two-step assembly protocol into two different G2-heart dimers (HA1-A1 and HA2-A2) with distinctive patterns of twelve units (Figures 1, 3, S8; Table S1D). AFM reveals that both samples exhibit a greater surface coverage than single G1 and G2-heart constructs (Figures 3, S14). It also demonstrates that single tail-tail interactions connecting two UAh_3WJ modules stack strongly enough to specify the directionality and orientation between two G2-hearts.

We hypothesized that G2-heart units could be encoded within a single-stranded RNA to generate the G3-heart, a ~715 nucleotides long molecule anticipated to fold during transcription into a heart shape. Encoded units should act as modular structural domains that promote folding under thermodynamic and kinetic control in isothermal conditions[29]. Depending on the way these domains are encoded within the RNA sequence, the location of kissing-loops, and choice of 5'-3' ends, it is theoretically possible to design multiple constructs with distinct folding pathways (Figures 4A, S15). Therefore, G3-heart constructs with efficient co-transcriptional folding are supposed to favor local formation of modular RNA subdomains while minimizing the length of transient single-stranded regions occurring during transcription. In the G2-heart, units HC, HE and HF, which contain the A-minor junction and AMP aptamer modules, are structurally more complex than units HA, HB and HD, which only relies on the UAh_3WJ module. Among a total of fourteen different constructs with their 5'-3' ends localized in domain dA, we chose the construct KL1-6, which promotes heart closure through KL1 and KL6 (Figures 4, S15; Table S1e). This construct maximizes the local modular formation of domains dC, dE and dF while also minimizing the length of transient single stranded regions during transcription. G3-heart co-transcriptional folding was investigated by AFM under solution[22] (Figures 4, S16). Approximately 70% of the objects on the mica surface correspond to G3-heart single particles. Multiple AFM scans of the same regions on the mica surface suggest that most G3-hearts are well formed with structural characteristics similar to G2-hearts (Figures 4C-F, S16). G3-hearts have the same preferential orientation as G2-hearts when lying on the mica surface. The degrees

of structural flexibility observed for few particles likely result from interacting with the scanning probe during AFM imaging under solution.

Drawing an analogy between the structural organization of RNA architectures and the way words are put together to form phrases, clauses or sentences in linguistics, we have demonstrated herein that a syntax of RNA 3D modules can be used to program RNA strands to fold into complex, responsive nanostructures with well-defined 3D shapes. While these structural modules can control the multimeric assembly of distinct RNA units, they can also be encoded within long unique RNAs to promote the isothermal folding and assembly of single-stranded RNA nano-architectures during transcription. The latter approach has great potential for low copy number expression of responsive RNA scaffolds inside the cell[30, 31]. Our RNA syntax was used herein to express a symbol of hope and love, but it can also be applied to the design of denser, solid 3D nanostructures conceivably as complex as the ribosome. In summary, this work emphasizes the remarkable amenability of RNA to rational design and the limitless potential of expression that RNA offers for engineering complex 3D nanostructures for synthetic biology, nanomedicine and nanotechnology[1, 3-6, 26]. By testing the orientational and topological behavior of RNA tertiary modules, our approach can significantly contribute to solving the sequence to 3D structure relationship in RNA folding. Finally, our "Heart" study suggests that stable RNAs, like the ribosome, might have increased in size and complexity during evolution by recombination and accretion of modular self-assembling domains at kissing-loop recombination sites[32, 33].

RNA 3D Modeling and Sequence Design. RNA design and nano-construction based on RNA structural modules (Figure S1) is described in the Supplementary Information (SI) and Figure S2. The heart-shaped atomic models were manually constructed "*in silico*" with the software SwissPdbviewer (http://spdbv.vital-it.ch/) by combining RNA fragments extracted from known x-ray crystallographic structures (http://www.rcsb.org). The merged RNA fragments were then refined to fix the backbone stereochemistry at the locations of the splices with Assemble 2.0 (http://bioinformatics.org/assemble/) and UCSF Chimera (http://www.cgl.ucsf.edu/chimera/). RNA sequences folding into a unique 2D structure were optimized in order to minimize alternative folds using NUPACK and the Vienna 2.0 package.

RNA synthesis. All RNA molecules were prepared by *in vitro* T7 RNA polymerase run-off transcription from PCR generated templates as previously described[2] . For native PAGE, TGGE and

structural probing, RNA assembly was typically monitored by PAGE using ^{32}P-alpha-ATP labeled RNA or 3'-[^{32}P]pCp labeled RNA as previously described[2] (see also SI).

RNA self-assembly, native PAGE and TGGE assays. *Folding Protocol 1:* RNAs are prepared by mixing different tectoRNAs (200 nM each, unless otherwise specified) in water. RNA samples were subjected to a denaturation/renaturating step by heating the samples to 90°C for 3 mins, cooling on ice 3 mins, incubation at 30°C for 3 mins, followed by addition of 5X concentrated buffer to reach 10 mM Tris-borate pH 8.2 (TB), 50 mM KCl, and 0.2 mM $Mg(OAc)_2$, prior further incubation at 30°C for 30 mins.

Folding protocol 2: in addition to protocol 1, RNA nanostructures were further stabilized for native PAGE and AFM studies, by raising the magnesium concentration to 15 mM $Mg(OAc)_2$ followed by heating at 60°C for 10 min and slow cooled to 10°C over 30 min.

Native PAGE and TGGE experiments were performed essentially as described in references[18, 23] (see SI for details).

Structural probing. RNAs were assembled *via* protocol 1. Chemical probing were performed at various concentrations of Pb(II) according to established protocols[18] (see SI for details). The RNA fragments were separated on denaturing polyacrylamide gels (15% acrylamide 29:1, 8 M urea) before quantitation using a phosphoimager.

AFM characterization of multimeric nanostructures. RNA architectures were assembled in solution, deposited on a mica surface in 15 mM Mg^{2+} buffer and visualized either under solution or dried under nitrogen before AFM imaging in air (3WJ-based heart, G2) (see SI for details). Multimode AFM imaging was performed at room temperature in tapping mode using a Nanoscope IIIa controller (Veeco, Santa Barbara).

G3-heart co-transcriptional folding and AFM visualization. RNAs were co-transcriptionally folded at 37.0°C for 20 minutes in one-pot reactions (50 µl) containing the template DNA (4 ng/µl final), NTPs (0.5 mM each final), DTT (1 mM final), 6 mM $Mg(OAc)_2$, 40 mM Na-OAc, 40 mM KCl, and 50 mM Tris-OAc (pH 7.8) in presence of T7 RNA polymerase (~0.2 U/50 µl). For AFM imaging, 5 µl of the transcription mix was directly mixed with AFM buffer (12.5 mM $Mg(OAc)_2$, 40 mM KCl, 40 mM NaCl, TB pH7.8) on freshly-cleaved mica surface, preheated to 37.0°C. After 1minute, the fluid was removed from the surface and addition of fresh AFM buffer (40 µl) and 4 µl of 60 mM $NiCl_2$. Multimode imaging was performed with 9.7khz "solution imaging" tips under these AFM conditions.

Cryo-EM for imaging G2-hearts. RNA heart samples were prepared according to protocol 2. Micrographs were acquired using a Tecnai F20 Twin transmission electron microscope operating at 120 kV, a nominal magnification of 80,000X, a pixel size of 0.105 nm at the specimen level, and a dose of ~30 e-/Å2. Images were automatically collected by the Leginon system[34] and recorded with a Tietz F415 4k x 4k pixel CCD camera. Experimental data were processed using the Appion software package[35, 36]. 4,467 particles were automatically selected from the micrographs using the FindEM, template-based particle picker[37]. The particle class averages were created using the Xmipp MLalign2D program[38] for reference-free alignment and classification using the fast algorithm[39]. A resolution of 20.8 Å for the class average based on 2,456 particles was determined by converting the spectral signal-to-noise ratio (SSNR) into a Fourier ring correlation (FRC) taken at a cutoff of 0.5.[40]

Supporting Information: Materials and Methods; Supplementary Tables S1A to S1E; Supplementary Figures S1−S16: Supplementary References.

Acknowledgements: LJ thank Maria del Carmen Jaeger for her silent contribution during the twelve years of this project. The authors thank E. Jacovetty, C. Potter and B. Carragher for their scientific input regarding cryo-EM and single-particle reconstruction. Cryo-EM was performed at the National Resource for Automated Molecular Microscopy, which is supported by the NIH National Institute of General Medical Sciences (GM103310). This work was supported by the National Institutes of Health (grants R01GM079604 and 5P41RR017573-10 to LJ), and intramural research grants from the UCSB Academic Senate (to LJ). CG acknowledges support from US National Science Foundation Expedition in Computing (1317694), and support from the Carlsberg Foundation. This paper is dedicated to the Three Hearts of the Holy Family in memory of Prof. Albert Jaeger, MD.

Author contributions: L.J. and C.G. conceived and designed experiments; C.G., A.C., E.V. performed biochemical experiments; A.C., C.G. performed AFM; N.V. performed cryo-EM: L.J. wrote the paper with the help of C.G. and A.C.

Competing financial interest:. There is no competing financial interest.

Materials & Correspondence*:* jaeger@chem.ucsb.edu; phone: 805 893 3628; Fax: 805 893 4120

References

(1) Jaeger, L.; Leontis, N. B. *Angew. Chemie. Int. Ed.* **2000,** 14, 2521-2524.
(2) Jaeger, L.; Westhof, E.; Leontis, N. B. *Nucleic Acids Res.* **2001,** 29, 455-63.
(3) Guo, P. *Nat. Nanotechnol.* **2010,** 5, 833-42.
(4) Grabow, W.; Jaeger, L. *F1000Prime Rep.* **2013,** 5, 46.
(5) Grabow, W. W.; Jaeger, L. *Acc. Chem. Res.* **2014,** 47, 1871-80.
(6) Afonin, K. A.; Kasprzak, W. K.; Bindewald, E.; Kireeva, M.; Viard, M.; Kashlev, M.; Shapiro, B. A. *Acc. Chem. Res.* **2014**.
(7) Yesselman, J. D.; Das, R. *Nucleic Acids Res.* **2015,** 43, W498-501.
(8)Anderson-Lee, J.; Fisker, E.; Kosaraju, V.; Wu, M.; Kong, J.; Lee, J.; Lee, M.; Zada, M.; Treuille, A.; Das, R. *J. Mol. Biol.* **2016,** 428, 748-57.
(9) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011,** 44, 1302-11.
(10) Jaeger, L.; Verzemnieks, E. J.; Geary, C. *Nucleic Acids Res.* **2009,** 37, 215--230.
(11) Apostolico, A.; Ciriello, G.; Guerra, C.; Heitsch, C. E.; Hsiao, C.; Williams, L. D. *Nucleic Acids Res.* **2009,** 37, e29.
(12) Gracia, B.; Xue, Y.; Bisaria, N.; Herschlag, D.; Al-Hashimi, H. M.; Russell, R. *Journal of molecular biology* **2016,** 428, 3972-3985.
(13) Westhof, E.; Masquida, B.; Jaeger, L. *Fold. and Des.* **1996,** 1, R78-R88.
(14) Saito, H.; Inoue, T. *Int. J. Biochem. Cell. Biol.* **2009,** 41, 398-404.
(15) Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. *Science* **2004,** 306, 2068-72.
(16) Grabow, W. W.; Zhuang, Z.; Swank, Z. N.; Shea, J. E.; Jaeger, L. *J. Mol. Biol.* **2012,** 424, 54-67.
(17) Afonin, K. A.; Bindewald, E.; Yaghoubian, A. J.; Voss, N.; Jacovetty, E.; Shapiro, B. A.; Jaeger, L. *Nat. Nanotechnol.* **2010,** 5, 676-82.
(18) Severcan, I.; Geary, C.; Chworos, A.; Voss, N.; Jacovetty, E.; Jaeger, L. *Nat. Chem.* **2010,** 2, 772-9.
(19) Geary, C.; Chworos, A.; Jaeger, L. *Nucleic Acids Res.* **2011,** 39, 1066
(20) Oi, H.; Fujita, D.; Suzuki, Y.; Sugiyama, H.; Endo, M.; Matsumura, S.; Ikawa, Y. *J. Biochem.* **2017,** 161, 451-462.
(21) Afonin, K. A.; Kireeva, M.; Grabow, W. W.; Kashlev, M.; Jaeger, L.; Shapiro, B. A. *Nano Lett.* **2012,** 12, 5192-5.
(22) Geary, C.; Rothemund, P. W.; Andersen, E. S. *Science* **2014,** 345, 799-804.
(23) Severcan, I.; Geary, C.; Verzemnieks, E.; Chworos, A.; Jaeger, L. *Nano Lett.* **2009,** 9, 1270-1277.
(24) Mohan, S.; Noller, H. F. *Nat. Commun.* **2017,** 8, 14285.
(25) Sassanfar, M.; Szostak, J. W. *Nature* **1993,** 364, 550-3.
(26) Dieckmann, T.; Butcher, S. E.; Sassanfar, M.; Szostak, J. W.; Feigon, J. *J. Mol. Biol.* **1997,** 273, 467-78.
(27) Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature* **1996,** 382, 183-6.
(28) Dibrov, S. M.; McLean, J.; Parsons, J.; Hermann, T. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** 108, 6405-8.
(29) Pan, T.; Sosnick, T. *Annu. Rev. Biophys. Biomol. Struct.* **2006,** 35, 161-75.
(30) Delebecque, C. J.; Silver, P. A.; Lindner, A. B. *Nat. Protoc.* **2012,** 7, 1797-807.
(31) Sachdeva, G.; Garg, A.; Godding, D.; Way, J. C.; Silver, P. A. *Nucleic acids Res.* **2014,** 42, 9493-503.

(32) Ennifar, E.; Walter, P.; Ehresmann, B.; Ehresmann, C.; Dumas, P. *Nat. Struct. Biol.* **2001,** 8, 1064-8.
(33) Petrov, A. S.; Gulen, B.; Norris, A. M.; Kovacs, N. A.; Bernier, C. R.; Lanier, K. A.; Fox, G. E.; Harvey, S. C.; Wartell, R. M.; Hud, N. V.; Williams, L. D. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** 112, 15396-401.
(34) Suloway, C.; Pulokas, J.; Fellmann, D.; Cheng, A.; Guerra, F.; Quispe, J.; Stagg, S.; Potter, C. S.; Carragher, B. *J. Struct. Biol.* **2005,** 151, 41-60.
(35) Lander, G. C.; Stagg, S. M.; Voss, N. R.; Cheng, A.; Fellmann, D.; Pulokas, J.; Yoshioka, C.; Irving, C.; Mulder, A.; Lau, P.-W.; Lyumkis, D.; Potter, C. S.; Carragher, B. *J. Struct. Biol.* **2009,** 166, 95-102.
(36) Voss, N. R.; Lyumkis, D.; Cheng, A.; Lau, P. W.; Mulder, A.; Lander, G. C.; Brignole, E. J.; Fellmann, D.; Irving, C.; Jacovetty, E. L.; Leung, A.; Pulokas, J.; Quispe, J. D.; Winkler, H.; Yoshioka, C.; Carragher, B.; Potter, C. S. *J. Struct. Biol.* **2010,** 169, 389-98.
(37) Roseman, A. M. *Ultramicroscopy* **2003,** 94, 225-36.
(38) Scheres, S. H. W.; Valle, M.; Carazo, J. M. *Bioinformatics* **2005,** 21 Suppl 2, ii243-4.
(39) Scheres, S. H. W.; Valle, M.; Grob, P.; Nogales, E.; Carazo, J. M. *J. Struct. Biol.* **2009,** 166, 234-40.
(40) Unser, M.; Trus, B. L.; Steven, A. C. *Ultramicroscopy* **1987,** 23, 39-51.

Figure Legends

Figure 1: The RNA structural syntax behind the rational design of 3D RNA nanostructures. (left) RNA 3D modules defining angles ranging from 60° to 180°; (middle) set of different RNA domains, or tectoRNA units, derived from the 3D modules. Arrows indicate the structural relationship between tectoRNAs; (right) Examples of RNA shapes described in the text. Color code for modules is the same for all panels. For more details, see also Supplementary Information Figures S1, S2.

Figure 2: RNA modules control the assembly of tectoRNA units into repetitious nanostructures. **A**. Assembly principle of "obtuse" or "acute" UAh_3WJ RNA units into tectosquares (TS and TSa), tectotriangles (TT and TTo) and hexameric nanorings (2TTo). Tectosquares formed of four units (A, B, C and D) assemble through four different kissing loops. Tectotriangles and hexamers are formed of three units (A', B and C'). Tm's obtained by TGGE at 0.2 mM $Mg(OAc)_2$ are indicated for each assembly (Materials and Methods). **B**. TS, TT and TTo constructs can assemble further through complementary 3' tails, to form programmable nano-arrays and nanogrids (Figure S3). Except for TT5+TT6+TT7 imaged in air, AFM visualization was performed under solution with 15mM $Mg(OAc)_2$, as described in the Material and Methods and Figure S5. All scale bars are 50 nm. **C**. Principle of assembly of AMP-TF units responsive to AMP, into nanorings and fibers of different size and length. **D.** Self-assembly of AMP-TF2 units in absence or presence of AMP at the indicated concentration is monitored by native PAGE at 1mM $Mg(OAc)_2$ and 10°C (Materials and Methods). AMP-TF2 variant is shown Figure S6. **E.** AFM images of AMP-TF1 nanofibers (25 nM) acquired under solution in presence of 2 mM ATP and 15 mM $Mg(OAc)_2$ at 20°C.

Figure 3: G1 and G2 nano-hearts have characteristic heart shapes responsive to AMP. **A.** G1-heart unit Y rigidified upon addition of 1mM AMP. Pb(II) (42 mM) cleavage profiles of HY within G1-hearts, with (red) or without (blue) AMP, compared to those of HY alone, with (green) or without (yellow) AMP (Figure S10). Boxed 3D model of HY: regions in yellow and blue have reduced Pb(II) cleavage with AMP and units H-WXZ; regions in red retain significant cleavage. **B.** AFM images of G1-Hearts obtained under magnesium solution (Figure S11). **C-D.** AFM images of G2-hearts acquired in air. **E.** Single G2-heart images. **F.** High resolution AFM image contour-plot of G2-heart. Each contour represents a 0.1 nm change in height. **G.** Same image compared to the G2-heart 3D model.

Tall points fit to tertiary modules. **H.** Cryo-EM image of G2-hearts on a thin layer of carbon. **I.** Combined left/right reference-based alignments of 2456 single G2-heart particles obtained from cryo-EM images. For (**E–I**), see also Figures S12, S13. **J.** G2-hearts with two different orientations of unit HA (HA1 or HA2) assemble into two distinctive dimers when linked through one tail-tail interaction. AFM images acquired in air (Figures S8, S14).

Figure 4: Single-stranded G3-hearts, with structural modularity similar to G2-hearts, are produced under kinetic control in isothermal conditions during transcription. **A.** Design criteria for the G3-heart KL1-6 (Materials and Methods; Figure S15). **B.** Co-transcriptional folding and visualization procedure for G3-hearts. **C.** Typical AFM images of G3-hearts acquired in solution according to procedure in (B). A portion of the mica surface displayed on the right was scanned twice by AFM. Most hearts do not change much in shape (circled in blue) but few can be disrupted by the AFM tip (circled in red). **D.** Yield of well-formed G3-hearts versus malformed or partial transcripts. **E.** Examples of well-resolved single G3-hearts from 0.5 μm AFM scans. **F.** Examples of well-resolved single G3-hearts from 1 μm AFM scans. For (C–F), see also Figure S16.

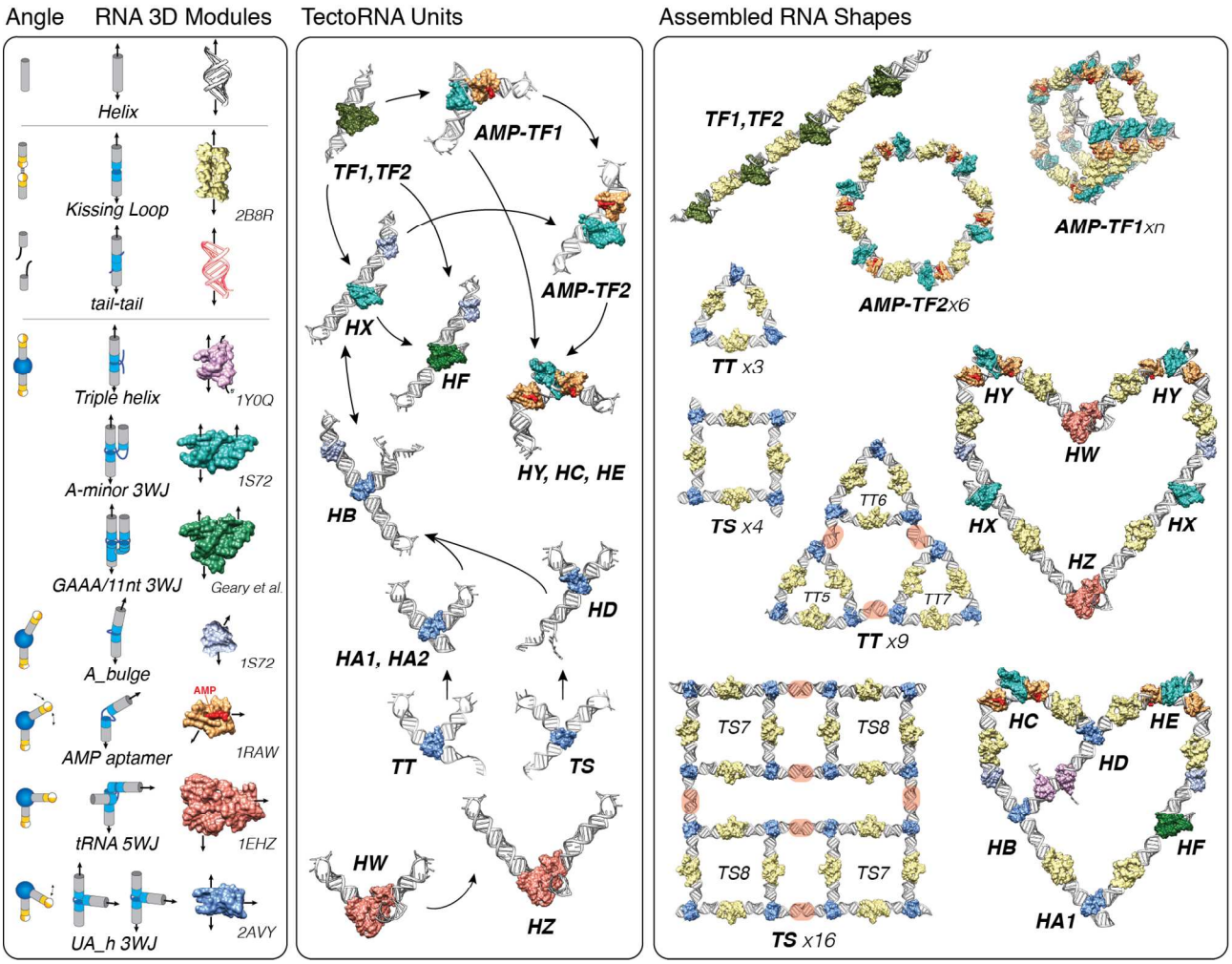

Angle
RNA 3D Modules
TectoRNA Units
Assembled RNA Shapes
Helix
Kissing Loop
2B8R
tail-tail
Triple helix
1Y0Q
A-minor 3WJ
1S72
GAAA/11nt 3WJ
Geary et al.
A_bulge
1S72
AMP
AMP aptamer
1RAW
tRNA 5WJ
1EHZ
UA_h 3WJ
2AVY
AMP-TF1
TF1,TF2
AMP-TF2
HX
HF
HY, HC, HE
HB
HD
HA1, HA2
TT
TS
HW
HZ
TF1,TF2
AMP-TF1xn
AMP-TF2x6
TT x3
HY
HY
HW
HX
HX
HZ
TS x4
TT6
TT5
TT7
TT x9
TS7
TS8
TS8
TS7
TS x16
HC
HE
HD
HB
HF
HA1

Figure 1

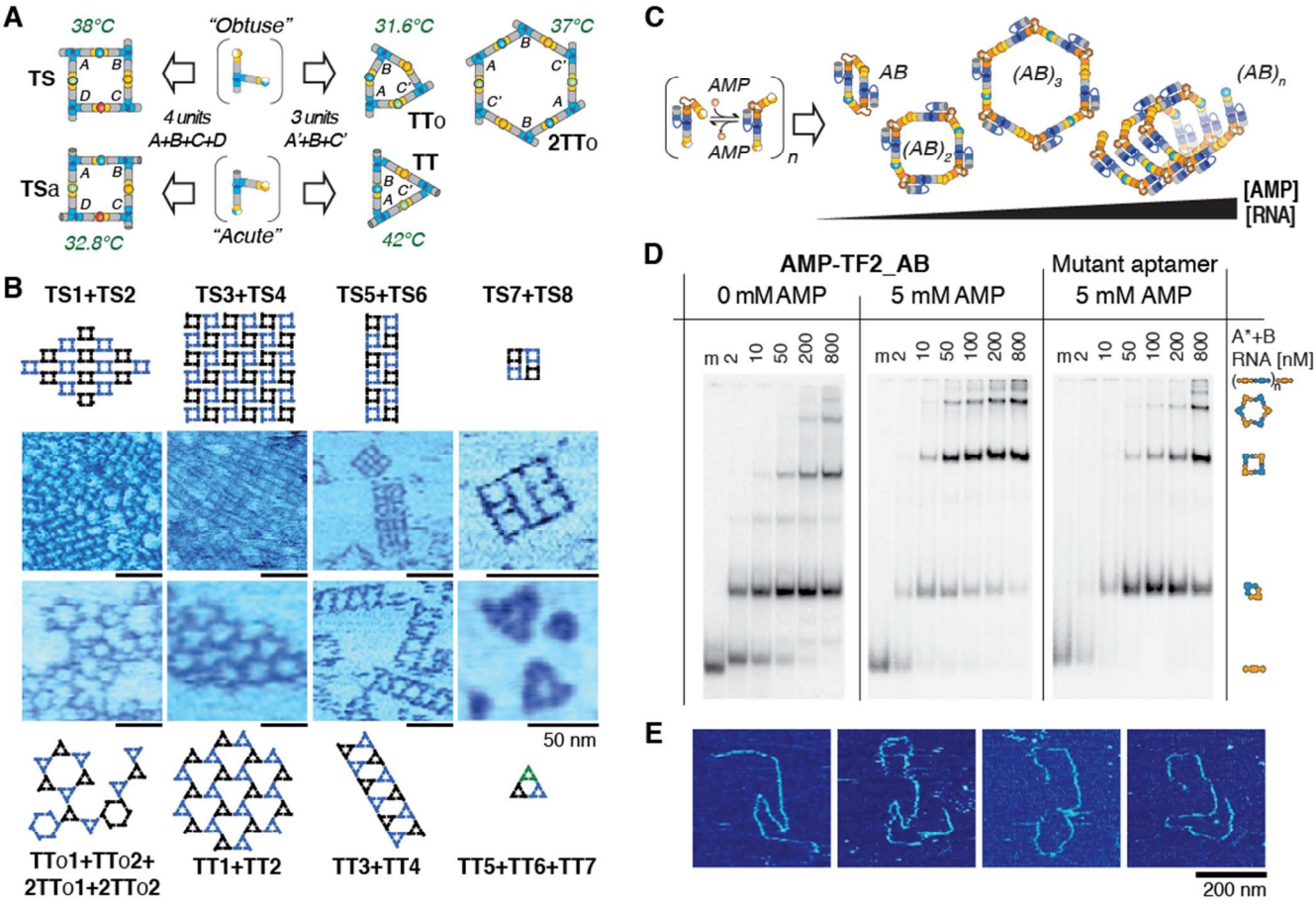

A
38°C
"Obtuse"
31.6°C
37°C
TS
4 units
A+B+C+D
3 units
A'+B+C'
TTo
2TTo
TSa
32.8°C
"Acute"
42°C
TT
B
TS1+TS2
TS3+TS4
TS5+TS6
TS7+TS8
50 nm
TTo1+TTo2+
2TTo1+2TTo2
TT1+TT2
TT3+TT4
TT5+TT6+TT7
C
AMP
AMP
n
AB
(AB)3
(AB)n
(AB)2
[AMP]
[RNA]
D
AMP-TF2_AB
0 mM AMP
5 mM AMP
Mutant aptamer
5 mM AMP
m 2 10 50 100 200 800
A*+B
RNA [nM]
E
200 nm

Figure 2

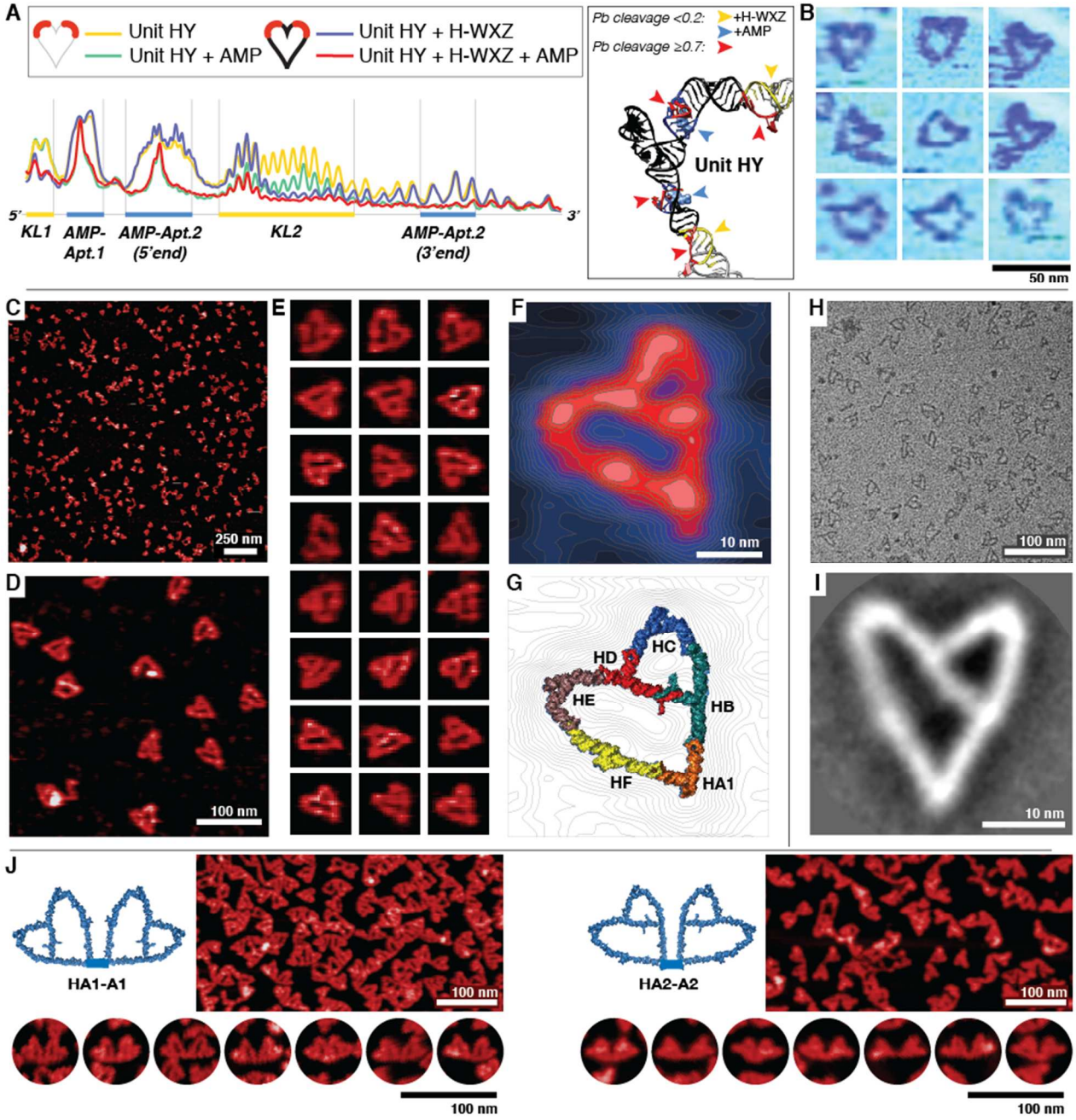

A
Unit HY
Unit HY + AMP
Unit HY + H-WXZ
Unit HY + H-WXZ + AMP
5'
3'
KL1
AMP-Apt.1
AMP-Apt.2 (5'end)
KL2
AMP-Apt.2 (3'end)
Pb cleavage <0.2:
+H-WXZ
Pb cleavage ≥0.7:
+AMP
Unit HY
B
50 nm
C
250 nm
D
100 nm
E
F
10 nm
G
HC
HD
HE
HB
HF
HA1
H
100 nm
I
10 nm
J
HA1-A1
100 nm
HA2-A2
100 nm
100 nm
100 nm

Figure 3

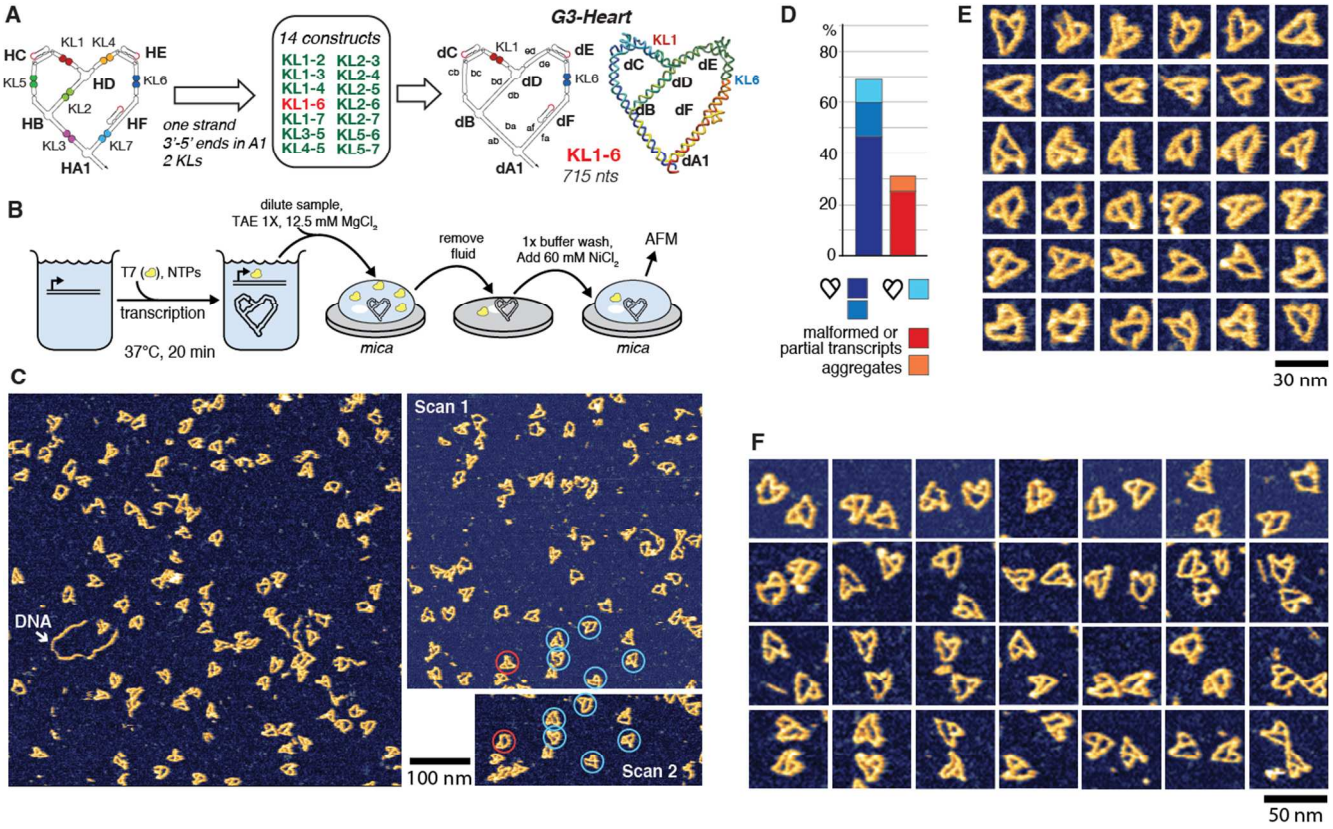
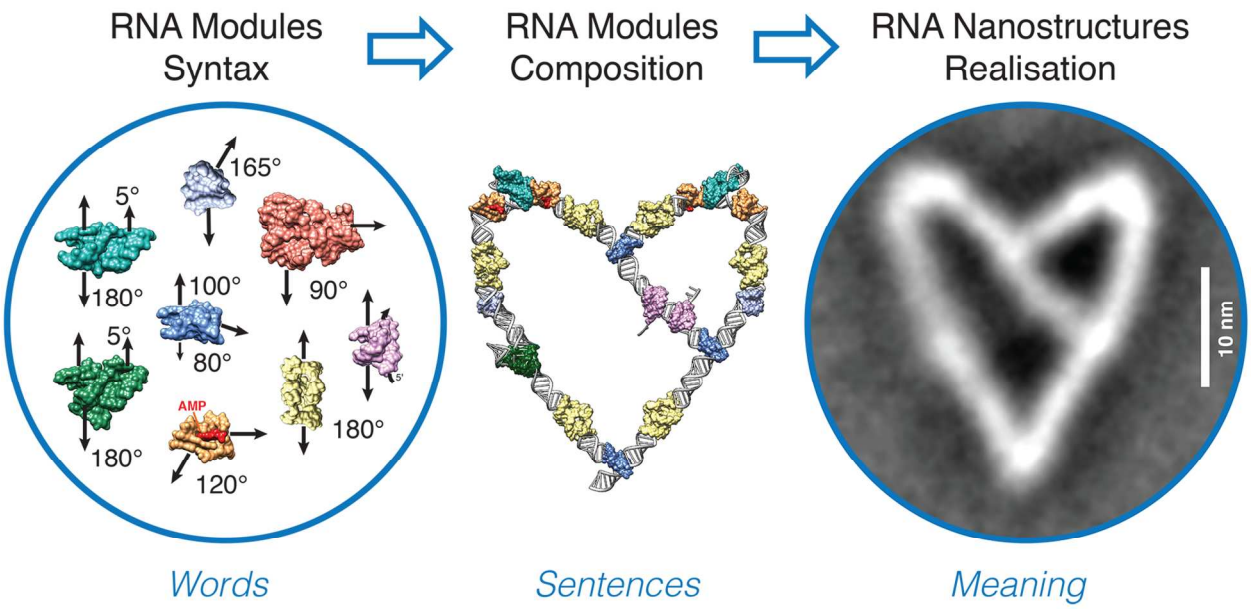

A
HC
KL1
KL4
HE
KL5
HD
KL6
KL2
HB
HF
KL3
KL7
HA1
one strand
3'-5' ends in A1
2 KLs
14 constructs
KL1-2 KL2-3
KL1-3 KL2-4
KL1-4 KL2-5
KL1-6 KL2-6
KL1-7 KL2-7
KL3-5 KL5-6
KL4-5 KL5-7
G3-Heart
dC
dE
dD
dB
dF
dA1
KL1-6
715 nts
B
dilute sample,
TAE 1X, 12.5 mM MgCl2
T7 (), NTPs
transcription
37°C, 20 min
mica
remove
fluid
1x buffer wash,
Add 60 mM NiCl2
AFM
mica
C
DNA
Scan 1
Scan 2
100 nm
D
%
80
60
40
20
0
malformed or
partial transcripts
aggregates
E
30 nm
F
50 nm

Figure 4

TOC graphic


RNA Modules Syntax
RNA Modules Composition
RNA Nanostructures Realisation
5°
165°
180°
100°
90°
5°
80°
AMP
180°
180°
120°
10 nm
Words
Sentences
Meaning